

## Research Paper

## Using GIS-linked Bayesian Belief Networks as a tool for modelling urban biodiversity



Darren R. Grafius<sup>a,b,\*</sup>, Ron Corstanje<sup>b</sup>, Philip H. Warren<sup>a</sup>, Karl L. Evans<sup>a</sup>, Briony A. Norton<sup>a</sup>, Gavin M. Siriwardena<sup>c</sup>, Oliver L. Pescott<sup>d</sup>, Kate E. Plummer<sup>c,e</sup>, Meghann Mears<sup>a</sup>, Joanna Zawadzka<sup>b</sup>, J. Paul Richards<sup>a</sup>, Jim A. Harris<sup>b</sup>

<sup>a</sup> University of Sheffield, Department of Animal and Plant Sciences, Sheffield S10 2TN, United Kingdom

<sup>b</sup> Cranfield University, School of Water, Energy and Environment, Cranfield, Bedfordshire MK43 0AL, United Kingdom

<sup>c</sup> British Trust for Ornithology, The Nunnery, Thetford, Norfolk IP24 2PU, United Kingdom

<sup>d</sup> Centre for Ecology and Hydrology, Wallingford OX10 8BB, United Kingdom

<sup>e</sup> Centre for Ecology and Conservation, College of Life and Environmental Sciences, University of Exeter, Penryn Campus, Cornwall TR10 9FE, United Kingdom

## ARTICLE INFO

## Keywords:

BBN  
Bird  
Invertebrate  
Fragmentation  
Model  
Species richness

## ABSTRACT

The ability to predict spatial variation in biodiversity is a long-standing but elusive objective of landscape ecology. It depends on a detailed understanding of relationships between landscape and patch structure and taxonomic richness, and accurate spatial modelling. Complex heterogeneous environments such as cities pose particular challenges, as well as heightened relevance, given the increasing rate of urbanisation globally. Here we use a GIS-linked Bayesian Belief Network approach to test whether landscape and patch structural characteristics (including vegetation height, green-space patch size and their connectivity) drive measured taxonomic richness of numerous invertebrate, plant, and avian groups. We find that modelled richness is typically higher in larger and better-connected green-spaces with taller vegetation, indicative of more complex vegetation structure and consistent with the principle of 'bigger, better, and more joined up'. Assessing the relative importance of these variables indicates that vegetation height is the most influential in determining richness for a majority of taxa. There is variation, however, between taxonomic groups in the relationships between richness and landscape structural characteristics, and the sensitivity of these relationships to particular predictors. Consequently, despite some broad commonalities, there will be trade-offs between different taxonomic groups when designing urban landscapes to maximise biodiversity. This research demonstrates the feasibility of using a GIS-coupled Bayesian Belief Network approach to model biodiversity at fine spatial scales in complex landscapes where current data and appropriate modelling approaches are lacking, and our findings have important implications for ecologists, conservationists and planners.

## 1. Introduction

A central and long-standing objective of landscape ecology is the ability to predict spatial variation in biodiversity. This requires accurate spatial biodiversity models at scales relevant to research and planning. Such tools support policies that aid conservation and optimise land-use patterns with minimal negative ecological impacts. They would also be valuable for assessing how land-use change influences ecosystem service provision, as biodiversity both underpins this provision (Millennium Ecosystem Assessment, 2005) and can itself be viewed as an ecosystem service (Bagstad, Semmens, Waage, & Winthrop, 2013).

Modelling biodiversity requires an understanding of how it responds

to landscape patterns, and which structural components are most influential in driving biodiversity responses. This can be challenging for numerous reasons. Aggregation of accumulated records at regional or national scales can characterise spatial patterns of biodiversity, at least for widely recorded groups (e.g. Anderson et al., 2009). It is extremely rare, however, for local scale biodiversity data on species occurrence or abundance to be available systematically across an entire study region (Gillespie et al., 2017). In such cases, there are generally three broad types of approach that can be used to model biodiversity across the region of interest. The first is modelling or extrapolating from species occurrence data, such as combining single species distribution models to predict spatial patterns of species richness across an area (e.g.

\* Corresponding author.

E-mail address: [d.grafius@sheffield.ac.uk](mailto:d.grafius@sheffield.ac.uk) (D.R. Grafius).

<https://doi.org/10.1016/j.landurbplan.2019.05.012>

Received 8 August 2018; Received in revised form 8 May 2019; Accepted 20 May 2019

0169-2046/ © 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Milanovich, Peterman, Barrett, & Hopton, 2012) or combining local species frequency curves to produce richness estimates for combined taxonomic groups (and to adjust for recorder effort; Dyer et al., 2017). Second, local-scale well-resolved biodiversity sample data can be combined with land cover types, climate and other environmental data to define the environmental 'envelope' relevant to certain taxa and then used to predict species presence/absence based on the occurrence or otherwise of suitable habitat, and this information can then be generalised over a wider study area and multiple species. Although informative, this approach lacks a directly modelled relationship between biodiversity and the environment (Massimino, Johnston, & Pearce-Higgins, 2015). The final approach uses habitat quality (e.g. a combination of area, habitat type, connectivity and threat) as a proxy for biodiversity or a measure of biodiversity potential (e.g. Kovacs et al., 2013). Biodiversity modelling for conservation planning often uses the first of these (Rodríguez, Brotons, Bustamante, & Seoane, 2007), whereas natural capital and ecosystem service planning uses the third (e.g. InVEST, LUCI; see Bagstad et al., 2013). Other methods such as hierarchical Bayesian approaches have also been explored to a lesser extent for hierarchical count data (Fordyce, Gompert, Forister, & Nice, 2011). These differences in approach tend to complicate the comparison and synthesis of model outcomes and agreement on optimum planning solutions. A key challenge remains how to predict landscape-scale spatial patterns in biodiversity by combining often-limited data (on both biodiversity and its predictors) with general ecological principles, so to generate predictions specific and sensitive enough to be useful for planning, conservation and natural capital assessment.

Urban areas provide an important situation in which the development of tractable biodiversity models is crucial. Urban development threatens some elements of biodiversity, yet urban areas often contain significant biodiversity, including threatened species (Ives et al., 2016). Towns and cities, at least in developed regions, are also highly regulated environments with a range of planning legislation and guidelines that create the potential to fine-tune detailed aspects of land use, including those affecting biodiversity (Norton, Evans, & Warren, 2016). Our ability to take advantage of this opportunity is limited, however, as current approaches to biodiversity modelling cannot generate accurate or useful predictions in urban landscapes due to high landscape complexity and heterogeneity (Kattwinkel, Strauss, Biedermann, & Kleyer, 2009). It is thus very difficult for planners to predict how biodiversity will respond to alternative urban designs.

Bayesian Belief Networks (BBNs), a form of probabilistic influence network, provide an alternative method with a number of potential advantages over the previously described approaches to biodiversity modelling that can be applied depending on the demands of the study. Bayesian approaches are historically varied and include methods that have a long history of use in text retrieval and medical diagnosis, where they are more commonly known as 'naïve Bayes classifiers' (Lewis, 1998; Pakhomov, Buntrock, & Chute, 2006). As networks, they can cope effectively with incomplete information on the relationships between variables, thus facilitating modelling when data availability is insufficient to render a deterministic approach feasible (Korb & Nicholson, 2010). BBNs are thus well-suited to complexity and incomplete knowledge, which are common in ecological systems (Jellinek, Rumpff, Driscoll, Parris, & Wintle, 2014; Landuyt et al., 2013). They can also be used to incorporate expert knowledge on the direction and strength of the effects of variables that influence biodiversity. Previously, BBNs could only operate on tabular data with model outcomes extracted and mapped separately (Aitkenhead & Aalders, 2009; Stassopoulou, Petrou, & Kittler, 1998). A recent advance, shared with other modelling approaches, is the ability to link BBNs with spatial data in order to incorporate spatial relationships into the model structure directly and to generate mapped model outcomes (e.g. predicted biodiversity) at a per-pixel level (Chee et al., 2016; Smith et al., 2018). This now allows BBNs to provide the mapped biodiversity predictions needed to inform planning and our understanding of biodiversity

responses to landscapes. Although linking BBN modelling with GIS functionality is not in itself novel, such methods have not previously been used to explore biodiversity in urban environments, which pose particular challenges for predicting biodiversity due to their highly heterogeneous and complex nature (Norton et al., 2016).

In this study, we present a GIS-linked BBN-based modelling method for predicting spatial patterns in urban biodiversity and use it to test relationships between urban habitat structure and biodiversity. We develop models based on mapped factors influencing biodiversity in urban areas described in a recent, and the first, meta-analysis of intra-urban biodiversity responses (Beninde, Veith, & Hochkirch, 2015). We then inform the BBNs with co-located richness data on plant, invertebrate and vertebrate (avian) communities sampled across three urban areas in southern England, and apply these spatially to produce biodiversity maps for entire urban areas at two different spatial resolutions to consider scale dependencies. The datasets used for predictors and, in particular, observed urban biodiversity represent reasonably complete and thorough data relative to those which are commonly available for urban areas. We assess the sensitivity of model output to the predictors in order to explore the relative influence of those drivers in determining predicted biodiversity. We also examine the correlations between taxa in their modelled spatial patterns of biodiversity to assess whether taxa are predicted to share common responses to the urban environment, and thus whether planning interventions are likely to benefit multiple taxonomic groups. Finally, we consider the key spatial patterns and drivers of urban biodiversity suggested by the model, and their consequences for urban conservation planning, and discuss the extent to which our modelling approach may be a practical tool for such work. This approach differs from past biodiversity modelling methods by fitting richness values directly to mapped predictors in a BBN modelling framework, with the key advantage of this approach being the ability to allow the model to generate its own conditional probabilities based on the available data. In interpreting model outcomes, we then focus on landscape and patch structural characteristics that are within the ability of urban planners and land managers to control. Our aim was to build on the current state of knowledge about the key drivers of urban biodiversity by testing consequent predictions using an approach capable of producing robust results at a sufficiently fine spatial scale to be relevant to urban planners and land managers.

## 2. Methods

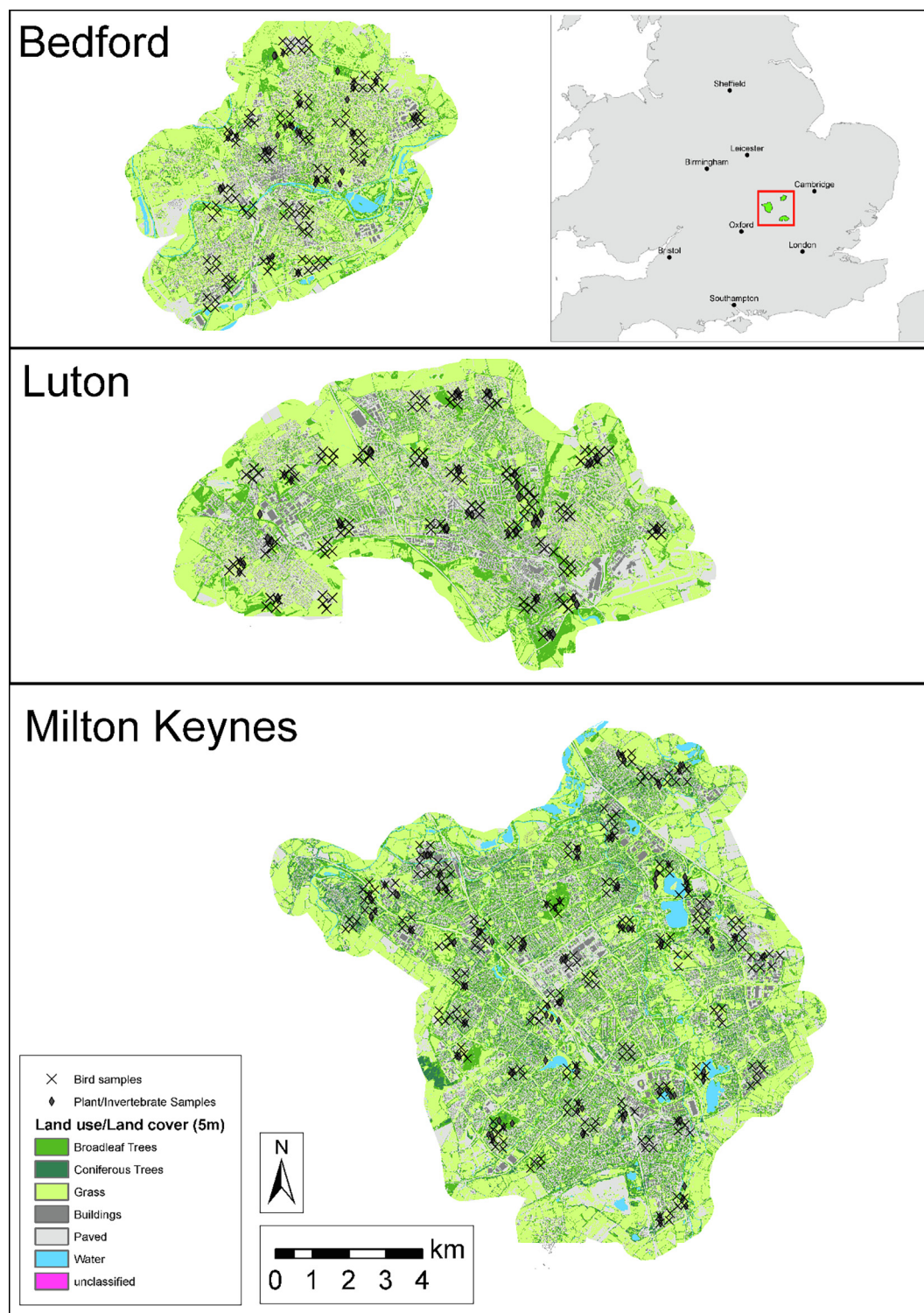
### 2.1. Study area

The study region was the combined built-up areas of three large UK towns that are close to each other but separated by areas of arable land, pasture, grassland and woodland (total urban area 183 km<sup>2</sup>: Milton Keynes, Bedford, and Luton; Fig. 1). The towns exhibit a broad range of urban forms and histories, capturing much of the diversity found across the UK's urban landscapes.

Milton Keynes (including Newport Pagnell and Bletchley; 52° 0' N, 0° 47' W) is a planned 'new town', developed during the 1960s. The town is structured around a grid of major roads designed for ease of automotive travel, and is characterised by large areas of public green space, consisting of parks and green areas bordering foot and cycle paths (Milton Keynes Council, 2015). Milton Keynes had a population of 229,941 in 2011, across 89 km<sup>2</sup> with a population density of 2584 inhabitants km<sup>-2</sup> (Office for National Statistics, 2013).

Bedford (52° 8' N, 0° 27' W) developed in the Middle Ages as a market centre and exhibits the radial development pattern typical to many British towns. In 2011, its population was 106,940 across 36 km<sup>2</sup>, with a population density of 2971 inhabitants km<sup>-2</sup> (Office for National Statistics, 2013).

Luton (Luton/Dunstable conurbation; 51° 52' N, 0° 25' W) developed heavily during the nineteenth century as an industrial centre. Its



**Fig. 1.** Study area locations and sampling locations for invertebrate/plant (diamonds) and bird (crosses) richness in Bedford, Luton and Milton Keynes, UK.

urban pattern contains large industrial zones and residential ‘terraced’ housing. The region had a 2011 population of 258,018 across 58 km<sup>2</sup>, with a population density of 4448 inhabitants km<sup>-2</sup> (Office for National Statistics, 2013).

## 2.2. Model formulation

Beninde et al. (2015) conducted a meta-analysis of 87 published studies on factors influencing urban biodiversity (as species richness or

species diversity, depending on the source) from 75 cities worldwide. This research looked at predictors including local (within-patch) and landscape (surrounding context) variables, as well as biotic (e.g. vegetation characteristics), abiotic (e.g. microclimate), and design (e.g. patch size) factors, each at an appropriate degree of specificity with respect to landscape and biodiversity measures (and the correlations between them). Species richness was found, in the majority of analysed papers, to exhibit positive relationships with patch size, corridor-based connectivity, and a range of vegetation factors including plant density,



vegetation structure and height and the amount of greenspace in the surrounding area.

We built our BBNs on the most feasible available representations of those predictors found to exhibit the strongest influences on urban biodiversity. These were implemented here as raster surfaces of patch size (area), corridor-based connectivity and vegetation height (see [Supplementary materials](#) for summary table). Vegetation height was used both due to its direct importance and as a composite proxy for influential vegetation structural factors, representing a single metric that is feasible to measure and interpret at landscape scales (described in more detail below). Another important identified predictor, amount of surrounding greenspace, was not addressed directly but is largely incorporated within our measures of connectivity and patch area.

### 2.3. Landscape factor input data: Patch area, connectivity and vegetation height

Patch area and connectivity data were based on a raster LULC map created from colour infrared aerial photography (<http://landmap.mimas.ac.uk/>). The imagery was taken on 2 June 2009 for Bedford, 30 June 2009 and 24 April 2010 for Luton, and 8 and 15 June 2007 and 2 June 2009 for Milton Keynes, based on cloud-free image availability. Vegetated and paved surfaces were separated using the Normalised Difference Vegetation Index (NDVI). UK MasterMap data ([Ordnance Survey \(GB\), 2017](#)) were subsequently used to identify buildings, water features, and major roadways. Habitat patches were defined as contiguous areas of greenspace regardless of vegetation type in order to maintain a generalised and multi-species perspective. Patch area (ha) was then calculated using Fragstats software ([McGarigal, Cushman, & Ene, 2012](#)) to encapsulate the amount of green space surrounding each point on the landscape. An 8-cell neighbourhood rule was used for patch inclusion, i.e. pixels were considered part of the same patch if they were diagonally as well as directly adjacent to one another. The resulting raster map of patch area (ha) was then incorporated a model predictor.

Corridor-based metrics of habitat connectivity have been identified as strong predictors of biodiversity in urban environments, whereas more generic connectivity metrics based on habitat proximity alone are less effective ([Beninde et al., 2015](#)). Circuit theory connectivity ([Dickson et al., 2018](#)) encapsulates this corridor-based approach and was thus used here, using Circuitscape software ([McRae, Shah, & Mohapatra, 2013](#)) to represent connectivity as a raster map of cumulative current, which is analogous to predicted wildlife flow or inverse cost-distance (see [McRae, Dickson, Keitt, & Shah, 2008](#)). Cumulative current data were obtained from Grafius et al. ([Grafius, Corstanje, Siriwardena, Plummer, & Harris, 2017](#)) who used the LULC map described above and data on Great Tit (*Parus major*) and Blue Tit (*Cyanistes caeruleus*) movements. These woodland species are sufficiently adaptive to be common in UK urban environments, but their movements within urban environments are constrained by habitat fragmentation ([Cox, Inger, Hancock, Anderson, & Gaston, 2016](#)) making them suitable species for corridor-based metrics of habitat connectivity. The connectivity metric is sensitive to the impacts of landscape features surrounding each pixel, and thus includes important elements of nearby landscape context in the models that are not accounted for by patch area or vegetation height.

Vegetation height (m) was measured using airborne LiDAR between June and September 2012 ([Casalegno, Anderson, Hancock, & Gaston, 2017](#); [Hancock, Anderson, Disney, & Gaston, 2017](#)) and used as a broad proxy for overall habitat maturity and vegetation structural diversity ([Bradbury et al., 2005](#)) in a form that was feasible to measure and to interpret at the landscape scale. Exact vegetation heights between 0 and 1 m, while expected to be ecologically meaningful, could not be reliably measured by the technology so we hereafter considered vegetated areas with heights of 0 m to represent mown lawn and 1 m to represent taller grasslands. The data were recorded at 2 m resolution, and aggregated

by mean value to 5 m resolution. Mean value aggregation was preferred to other methods as it preserves overall patch character, which we expected urban organisms to be more responsive to than patch extremes given the high heterogeneity of urban landscapes at fine scales ([Norton et al., 2016](#)). All predictor datasets were thus available at 5 m resolution and, additionally, following mean value aggregation, at 25 m resolution. The coarser scale was explored because it is comparable to the scale of widely available datasets such as the 25 m UK Land Cover Map 2015 and 30 m Landsat-derived land cover maps.

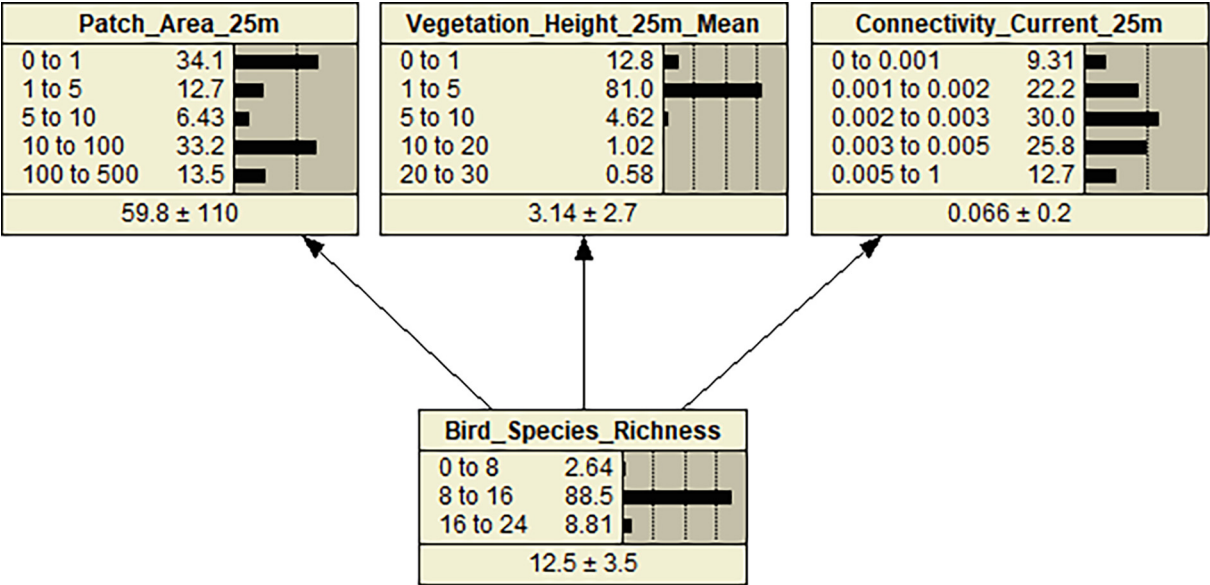
### 2.4. Biodiversity input data

Sampling across the study area using a stratified random design yielded datasets on the richness of eight taxonomic groups: invertebrates (to order level), litter organisms (to species level), Coleoptera (to family level), Diptera (to family level), birds (to species level) and non-tree plants (to species level, including all vascular plants from the ground, field and shrub layers). Plants were considered as total non-tree plants, as well as being separately categorised into native and neophyte non-tree plants for analysis, given the particular abundance of non-native plant species present in urban areas via landscaping, the expectation that they may be more strongly influenced by human activity and their available taxonomic resolution (to species level, as above). All the selected taxa are significant components of urban biodiversity and were selected because collectively they represent a wide range of habitat requirements, movement scales and ecological functions, and in some cases are particularly prominent forms of biodiversity that are important to human experience of nature (see [Supplementary Materials](#) for full details of sampling methodologies). Our primary interest in this study was the relative taxonomic richness. We use the term 'richness' to refer to the measured and modelled richness of each taxonomic group in accordance with the level of specificity described here. Note that, although data were gathered at three different time points (2009/2010 for aerial imagery, 2012 for LiDAR and 2013/2014 for biodiversity surveys), negligible change to urban land cover in the study area took place during this time.

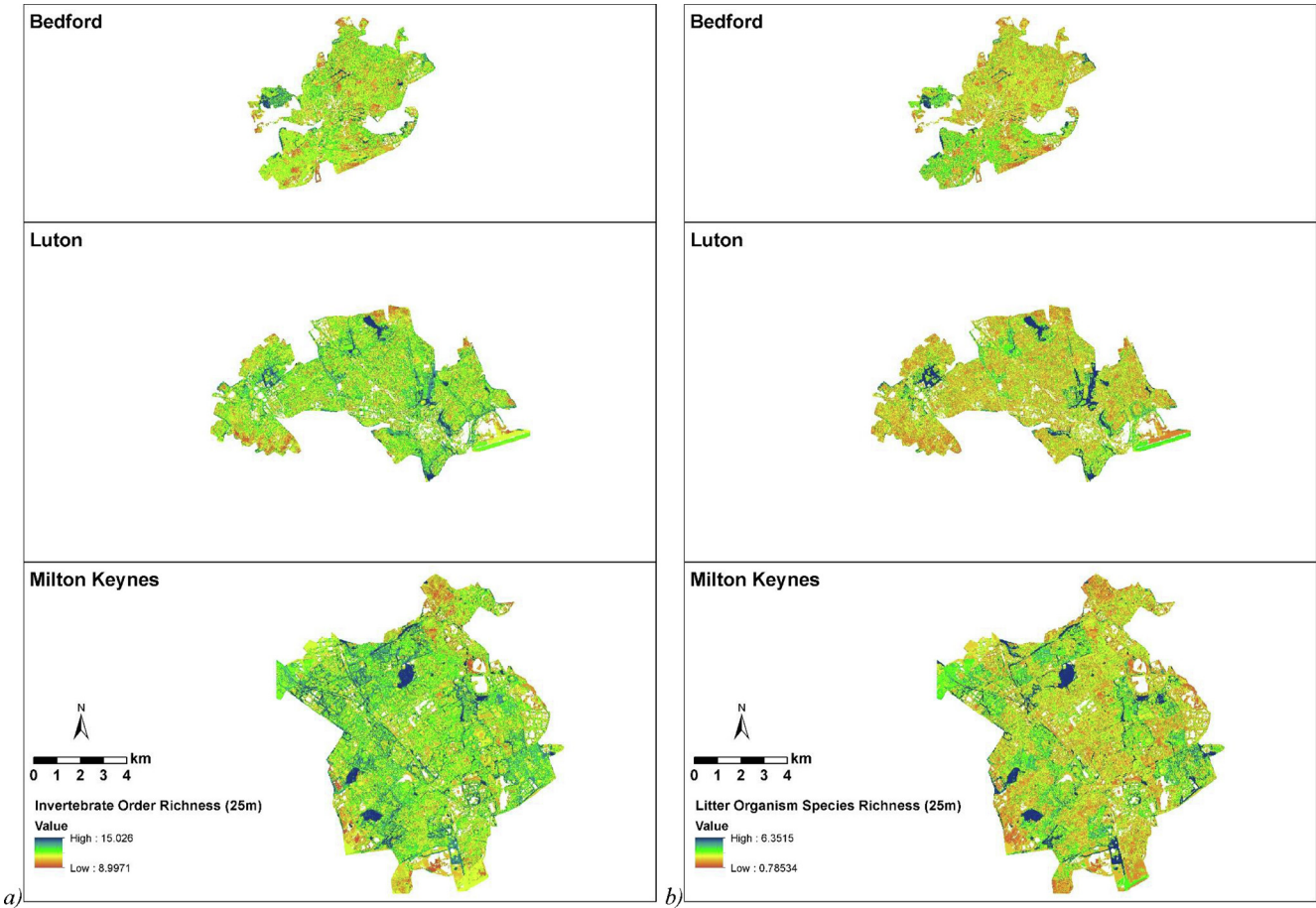
### 2.5. Model construction

BBN modelling was conducted using Netica software ([Norsys, 2016](#)) with the 'GeoNetica' extension enabling the integration of predictor datasets directly as spatial data. The BBN made predictions within vegetated areas, as most biodiversity observations and predictions were limited to these areas. Bird surveys were centred on vegetated areas, but encompassed mixed areas containing both vegetated and non-vegetated surfaces, e.g. paved surfaces, buildings and water.

Separate but comparable (i.e. possessing the same network structure) BBN models were created for each of the nine taxonomic groups, with each group's taxonomic richness as the model outcome. The influence network for each BBN included patch area, connectivity and vegetation height as model predictors, and biodiversity (richness) as the response variable. The predictor variables were chosen for their theorised direct influence on urban biodiversity; a decision that represents the primary input of expert knowledge into our modelling process (as opposed to the direct definition of conditional probabilities, which, as described previously, is also a possibility in BBN modelling where data availability is poorer). Conditional probabilities define the relationships between landscape factors and biodiversity, and were obtained through processing individual 'cases', where each case represented a point observation of richness and landscape factor values found at the same location (mean vegetation height within the 200 m radius used for birds). The model then used these conditional probabilities to predict taxonomic richness at every vegetated location within the study area. All nodes were automatically discretised with ten states each based on histogram equalisation for primary modelling. A simplification of this with five states for each input parameter and three



**Fig. 2.** Example Bayesian Belief Network model structure for Bird Species Richness. All models used a comparable structure, with only the dependent variable (i.e. taxonomic richness) and conditional probabilities changing between models. Arrows denote the direction of probabilistic influence implemented in software rather than causal relationships between the factors.



**Fig. 3.** Predicted taxonomic richness (maximum probability value/mode) at 25 m resolution for Bedford, Luton and Milton Keynes, UK: a) invertebrate order-level richness, b) species-level litter organism (Isopoda, Diplopoda, Chilopoda, and Pseudoscorpiones) richness, c) family-level Coleoptera richness, d) family-level Diptera richness, e) total non-tree plant species richness, f) native non-tree plant species richness, g) neophyte non-tree plant species richness, and h) bird species richness.

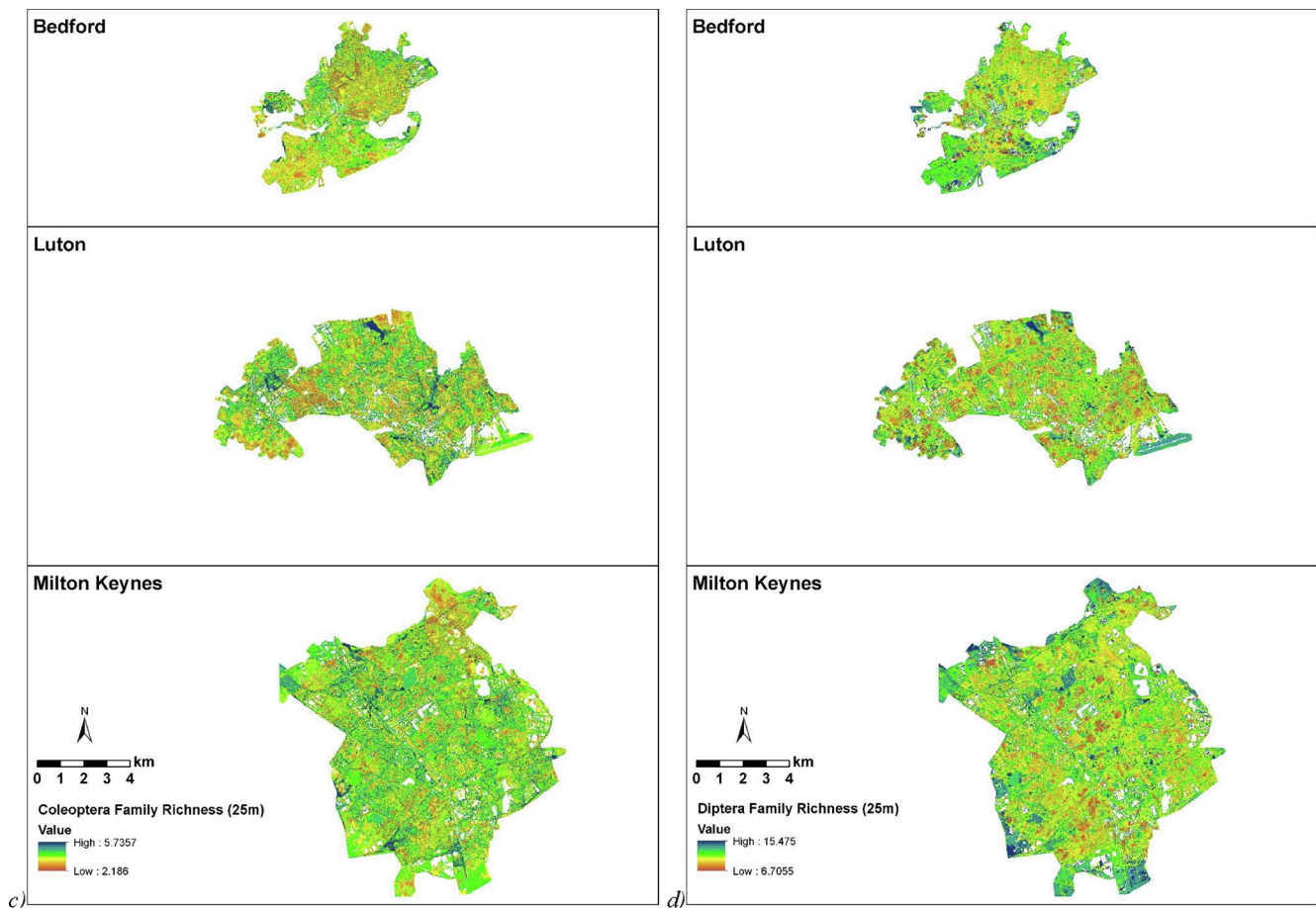


Fig. 3. (continued)

for dependent variables was used for ease of visualisation of model structure (Fig. 2) and consistency when comparing conditional probabilities (discussed below).

## 2.6. Model performance and sensitivity testing

Model performance was assessed using a goodness-of-fit measure, reported as the model's error rate, that expresses the frequency with which the model's strongest prediction (most likely outcome) is incorrect against the observed data. It does not represent validation in the same sense as in deterministic modelling, due to the Bayesian probabilistic inference base, but supplies an analogous measure of confidence in the model predictions (Aalders, 2008; Taalab et al., 2015b). Sensitivity analysis determined how much the beliefs (i.e. biodiversity predictions) were influenced by each new finding in the predictor nodes (i.e. changes in patch area, connectivity and/or vegetation height). Sensitivity was expressed as the expected reduction in variance of the expected real value due to a finding in a particular node (e.g. complete insensitivity would occur if the addition of new data records to the existing model caused no reduction in this variance). The conditional probabilities for the node states were extracted from the models and graphed as a heat map to show the predicted factor probability at each state level of biodiversity (after Fraser et al. (2016)).

## 2.7. Correlations between predicted biodiversity metrics

Regularly-spaced point samples from the model output maps were tested for correlations between different models to quantify the degree of difference and similarity between model predictions for different taxa. A regular grid of points (spaced at 25 m intervals) was generated

and clipped to the map extents, then used to extract raster values for import into R where correlation tests were run (R Development Core Team, 2016). Spearman's rank correlation was used to evaluate agreement between the predicted biodiversity maps. The permutation-framework Spearman's method using the R 'coin' package (Hothorn, Hornik, van de Wiel, & Zeileis, 2006) was used to generate a reliable p-value statistic given large numbers of tied values in the data. Our objective here is to assess spatial patterns in the co-occurrence of low or high biodiversity values across different taxa, i.e. whether there is spatial congruence in the biodiversity 'hotspots' across taxa. We thus do not take spatial autocorrelation into account, as this would distort the spatial patterns of interest.

## 3. Results

### 3.1. Predicted richness

Output maps show predicted richness across the study area for total invertebrates, species-level litter organisms, family-level Coleoptera, family-level Diptera, total non-tree plant species, native non-tree plant species, neophyte non-tree plant species and bird species, according to 25 m resolution data (Fig. 3) and 5 m resolution data (Supplementary Materials). The BBN models calculate a probability range at each pixel rather than a single value; displayed maps depict the maximum probability value (mode) at the centre of this probability range. Summary statistics are given for sampled richness data in Table 1, and of prediction map values in Table 2 (for 5 m resolution results and mapped landscape factor inputs, see Supplementary Materials).



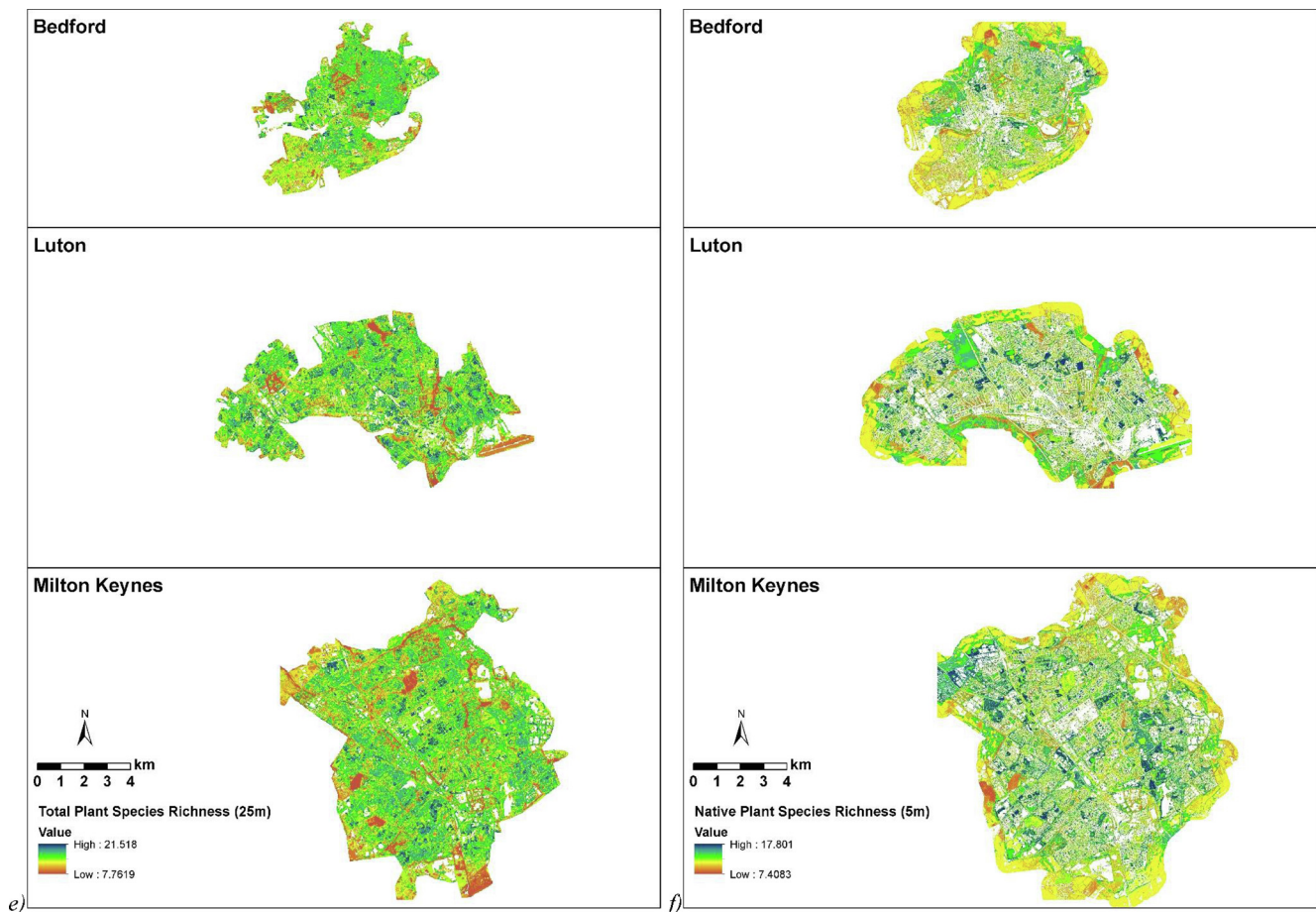


Fig. 3. (continued)

### 3.2. BBN model performance and predictor sensitivity

Error rates for 25 m models ranged between 46% for bird species richness and 69% for invertebrate order richness (Table 3), whereas the 5 m models exhibited error rates between 45% for neophyte plant richness and 74% for bird richness (see [Supplementary Materials](#)). Models at the two different spatial resolutions performed with largely comparable error rates, but we focus here primarily on the 25 m scale, given its comparability with available datasets and relevant scales of inquiry to ecologists and planners seeking to understand landscape-scale urban biodiversity.

Parameter sensitivities in the models reflect the strength of relationships between predictors and biodiversity predictions for each taxon (Table 3). For the 25 m models, parameter sensitivities were variable between taxa but vegetation height exhibited the greatest sensitivity for five of eight taxa. It also had the highest mean (3.3) and maximum (6.3) percentage variance reduction across all taxa combined (cf. connectivity, mean 2.4 and maximum 3.1; and patch area, mean 2.0 and maximum 3.1).

Predicted richnesses of total invertebrates, litter organisms, total plants, native plants and neophyte plants were most sensitive to vegetation height and exhibited relatively high percentage variance reduction (respectively 6.3, 5.3, 3.9, 3.2 and 3.2). Coleoptera richness predictions were most sensitive to connectivity but by a relatively small margin compared to patch area and vegetation height. Diptera and bird species richnesses were both most sensitive to patch area.

All measured plant richness predictions (total, native and neophyte) were most sensitive to vegetation height and least sensitive to patch area. The error rate for neophyte plant richness predictions was notably lower than other plant types, suggesting a stronger predictive ability.

Predicted bird richness was more sensitive to patch area than to other parameters (2.5), with responses to vegetation height exhibiting the lowest sensitivity (0.9).

The 5 m and 25 m models exhibited similar parameter sensitivities for most outcome variables but not all, e.g. Coleoptera was most sensitive to connectivity in the 25 m model and vegetation height in the 5 m model. Diptera was most sensitive to patch area at 25 m but vegetation height at 5 m, and native plants were most sensitive to patch area at 25 m but vegetation height at 5 m. The magnitude of scale dependence in the relative importance of key drivers of biodiversity may therefore vary by taxon (see [Supplementary Materials](#) for 5 m model results and performance).

### 3.3. Probabilistic associations between landscape factors and taxonomic richness

Heat maps show the nature of probabilistic associations between values of landscape factors and predicted richness levels using simplified node levels (Fig. 4, Table 4). Here, high conditional probabilities reflect the likelihood of an outcome given a set of parent node states; e.g. low Coleoptera family richness is expected in areas with low vegetation height, whereas high Coleoptera richness is expected in patches with moderately large area and high connectivity. Vegetation height appeared to be a strong predictor of total invertebrate richness, with the highest conditional probabilities associated with low to moderate vegetation height at moderate to high levels of predicted richness. Predicted litter organism richness exhibited strong associations with low vegetation height at low diversity levels, with patch area and connectivity appearing to play stronger roles at higher richness levels. Low levels of predicted Coleoptera and Diptera richness appeared to be

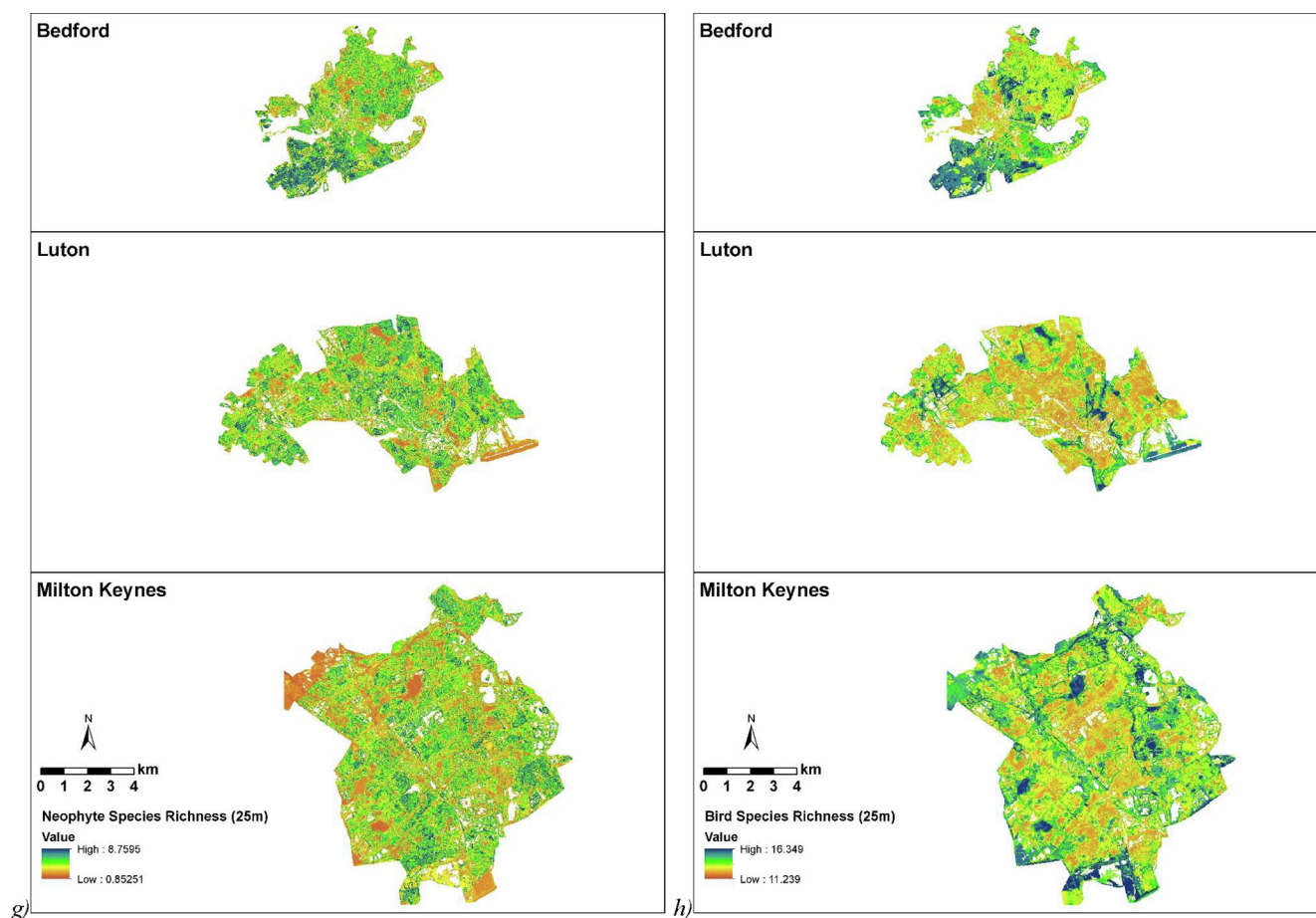


Fig. 3. (continued)

Table 1

Summary statistics of observations on invertebrate, plant and bird richness. 'Litter organisms' include species from Isopoda, Diplopoda, Chilopoda, and Pseudoscorpiones. Invertebrate samples were taken at the same sites as plant samples, but only contain  $n = 243$  observations due to a sweep net malfunction at one site.

	Mean	Std. Dev.	Min.	Max.	<i>n</i>
Invertebrate Order Richness	12.1	2.4	7	18	243
Litter Organisms Species Richness	2.1	2.6	0	14	243
Coleoptera Family Richness	4.0	2.3	0	12	243
Diptera Family Richness	9.5	4.7	0	25	243
Total Plant Species Richness	14.5	7.9	0	46	244
Native Plant Species Richness	11.0	7.9	0	46	244
Neophyte Plant Species Richness	2.9	6.0	0	30	244
Bird Species Richness	12.5	2.6	4	24	451

Table 2

Summary statistics of model prediction raster maps (25 m resolution) for invertebrate, plant and bird richness. 'Litter organisms' include species from Isopoda, Diplopoda, Chilopoda, and Pseudoscorpiones. Values are based on mean model results at each pixel.

	Mean	Std. Dev.	Min.	Max.
Invertebrate Order Richness	11.6	0.8	9	15
Litter Organisms Species Richness	2.0	0.8	1	6
Coleoptera Family Richness	4.0	0.6	2	6
Diptera Family Richness	10.0	1.3	7	15
Total Plant Species Richness	15.9	2.4	8	22
Native Plant Species Richness	12.1	1.8	7	19
Neophyte Plant Species Richness	3.6	1.5	1	9
Bird Species Richness	12.9	0.8	11	16

associated with moderately low vegetation heights. High total and neophyte plant richness predictions were associated with moderately low vegetation heights and moderately high connectivity. Lastly, all levels of predicted bird richness (i.e. low as well as high richness) showed strong associations with moderately low vegetation heights (1–5 m, i.e. shrubs and small trees).

### 3.4. Correlations between model predictions

High variation was present in the size and direction of correlation coefficients between models for different taxa (Table 5). Positive and negative correlations were approximately equal in number, with positive correlations relatively weak in many cases (8 of 28 correlation coefficients were positive and greater than 0.15). Notably high positive correlations (defined here as  $> 0.5$ ) were found between total invertebrate richness and litter organism species richness (0.79), Coleoptera and Diptera richness (0.53), and between total and neophyte plant richness (0.56; note that neophyte and native plants are both included in total plant richness, so some self-correlation is expected). Negative correlations were generally weaker than the strongest positive correlations, with the strongest being between Diptera and total plant richness ( $-0.65$ ), bird and neophyte plant richness ( $-0.47$ ), and between bird and total plant species richness ( $-0.44$ ). There was also a notable negative correlation between neophyte and native plant richness ( $-0.25$ ).

## 4. Discussion and conclusions

The modelling approach described here represents a method for: (1) applying current knowledge about the factors, supported by evidence,



**Table 3**

Results of case testing (error rate) and sensitivity analysis (percent variance reduction as a metric of the relative importance of each input variable) on Bayesian Belief Network models for invertebrate, litter organism, Coleoptera, Diptera, plant and bird richness at 25 m resolution. For each taxonomic group the landscape factor showing the greatest sensitivity is shown in bold.

	Error rate	Sensitivity (Percent Variance Reduction)		
		<i>Vegetation Height</i>	<i>Connectivity</i>	<i>Patch Area</i>
Invertebrate Order Richness	69.26%	<b>6.32</b>	2.81	1.39
Litter Organisms Sp. Richness	54.51%	<b>5.30</b>	2.80	2.02
Coleoptera Family Richness	67.62%	1.81	<b>2.33</b>	2.03
Diptera Family Richness	64.75%	1.90	2.74	<b>3.05</b>
Total Plant Sp. Richness	65.88%	<b>3.87</b>	3.08	1.65
Native Plant Sp. Richness	67.62%	<b>3.17</b>	2.76	2.36
Neophyte Plant Sp. Richness	48.36%	<b>3.23</b>	1.05	0.70
Bird Sp. Richness	45.68%	0.86	1.37	<b>2.50</b>

that influence urban biodiversity; (2) considering these factors in the context of specific local data for the system or area being studied; (3) combining these in an influence network to depict spatial patterns of modelled biodiversity across entire urban areas at a sufficiently fine resolution to be responsive to variation in urban landscape form, and thereby relevant to planning considerations and landscape-scale research. The scientific basis for our model structure stems from the findings of the meta-analysis of Beninde et al. (2015); although meta-analyses have limitations, they provide a summary of our knowledge of a system and can form a suitable evidence base for defining the network of important influences on a variable of interest, which in turn can form the basis for modelling that system (Stewart, Higgins, Schünemann, & Meader, 2015). Both the 5 m and 25 m analyses produced largely comparable results and error rates, leading us to believe that both scales have value and may be worthy of further exploration. However, we feel that the 25 m analysis is likely to be of greatest relevance to urban planners and landscape managers given its comparability with available datasets and common scales of inquiry. We next consider the ecological implications of the model predictions (i.e. test results), the performance of the models, and the practicality and application of using this approach for urban conservation and planning.

#### 4.1. Predicted spatial patterns and key drivers of urban biodiversity

Broad generalities can be discerned from the prediction maps (Fig. 3) and the conditional probability heat maps (Fig. 4) about the relationships between urban biodiversity and landscape structure. Vegetation height emerged as an important factor for multiple taxa, with low or moderate heights associated with low biodiversity in litter organism, Coleoptera, Diptera and native plant richness. There is therefore an expectation that in areas dominated by low vegetation height (such as mown lawns), the biodiversity of these groups will be relatively low. Large woodlands (e.g. greater than 10 ha in size), which are generally high in all three landscape factors (see [Supplementary Materials](#) for landscape factor maps), by contrast, exhibited high predicted biodiversity for some taxa including invertebrates, litter organisms and birds (Fig. 3). Conditional probabilities (Fig. 4) further support the expectation that total invertebrate, litter organism, Coleoptera, Diptera and bird richness will be greater in larger patches, which is consistent with past research on species-area relationships in urban landscapes (Nielsen, van den Bosch, Maruthaveeran, & van den Bosch, 2014).

Taller vegetation is not, however, invariably associated with greater biodiversity. Richness predictions for all three plant types were generally low in areas of tall vegetation such as woodlands, and Diptera varied by scale. Since modelled plant taxa did not include trees, this suggests that urban ground, field and shrub layer plants are more diverse in areas of low to moderate vegetation height, and more constrained in woodlands. Conditional probabilities show that an intermediate level of vegetation height (1–5 m; second box from the left in

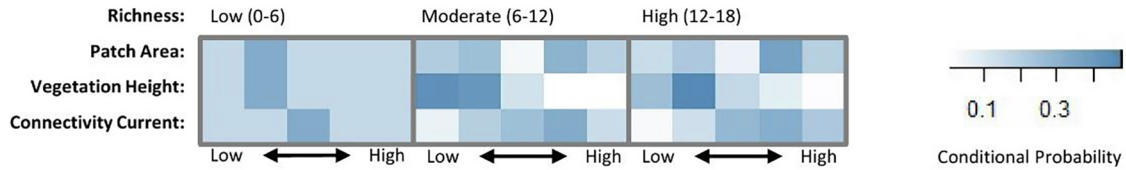
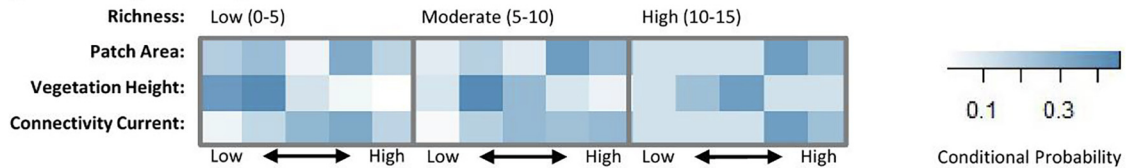
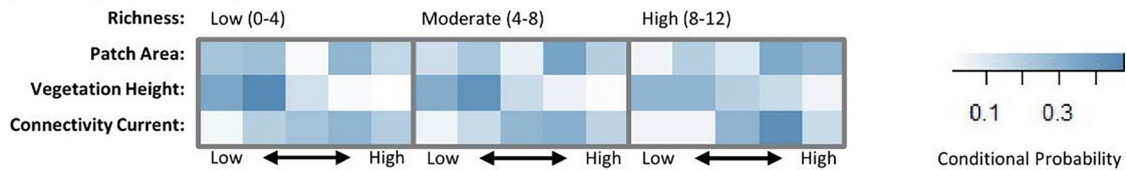
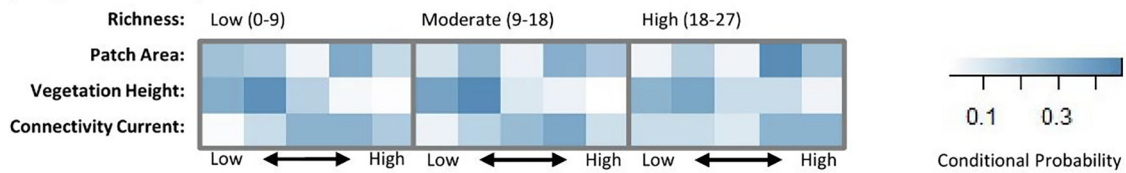
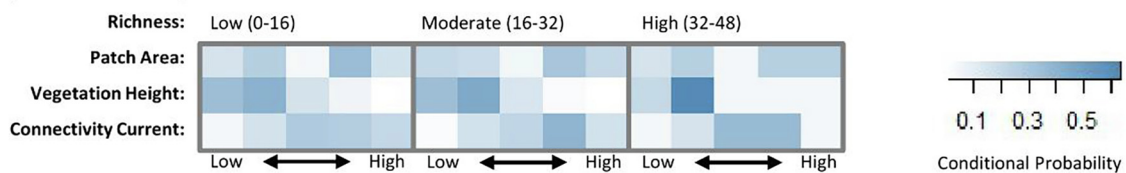
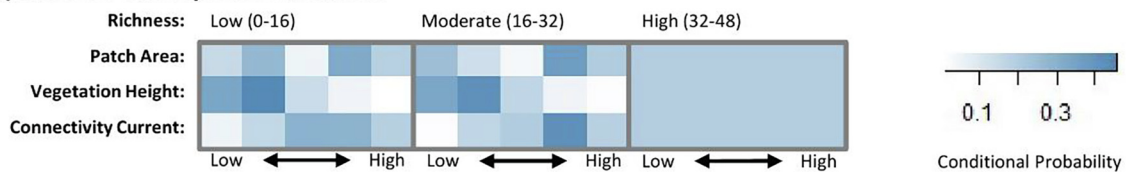
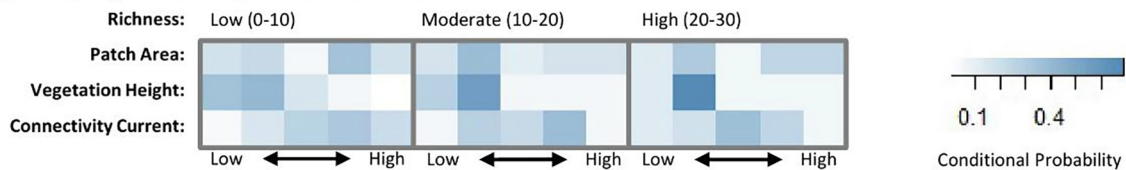
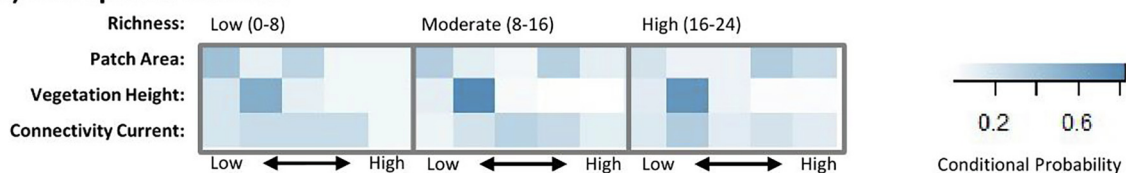
Fig. 4) is an important driver of richness predictions for invertebrates, total plant species, neophyte plant species and bird species, although the direction of association is not consistent across all taxa. Intermediate vegetation height was associated with high richness predictions for invertebrates, total and neophyte plants and birds, suggesting that meadows, hedges and shrubs at this height support biodiversity for these taxa. By contrast, litter organisms, Coleoptera, Diptera and native plants had lower richness predictions associated with these areas. The association between neophyte plants and intermediate vegetation height in particular may be consistent with the fact that some proportion of these species are likely to have been planted, for example in back gardens, rather than naturalised, particularly in urban settings. Additionally, many neophytes are ruderal species, well-adapted to colonising fragmented or more frequently disturbed habitats (Crawley, Harvey, & Purvis, 1996; Thompson & McCarthy, 2008) and are thus common in these urban areas of mixed vegetation.

Bird species richness did not show strong spatial relationships with input variables compared to other taxa. Predicted bird richness was most sensitive to patch area, which may be a consequence of the mobile nature of birds and associated larger home range sizes compared to our other focal groups. This suggests that habitat extent is a key driver of urban avian biodiversity, but the complexity of the response suggests that other factors are also important (Deák, Hüse, & Tóthmérész, 2016). The typically lower sensitivities of avian richness models relative to other taxa may partly arise from a mismatch in the scale of modelling compared to the larger home range size of most bird species, meaning that individual birds may use multiple, diverse habitat patches. In addition, a stronger effect of vegetation height and connectivity was expected (Hinsley et al., 2009) but its lack may be due to the absence of woodland specialists that are sensitive to fragmentation from urban areas in the UK. Whilst the urban avifauna mostly comprises species that are generalists (Evans, Chamberlain, Hatchwell, Gregory, & Gaston, 2011) individual species exhibit preference for a range of vegetation structures, from short open vegetation to tall trees, and thus summation of richness across all these types of species may further weaken relationships between species richness and patch size and vegetation height.

#### 4.2. Correlations between taxa

Managing urban areas to optimise biodiversity would be simpler if areas that supported high diversity of one group also supported high diversity of other groups. We found limited evidence for this as a slim majority of correlation coefficients were positive (15 compared to 13) and those that were positive were often rather limited in strength (12 out of 15 were lower than 0.5).

The predicted richness values of some taxa exhibited strong positive correlations with each other, suggesting possible similarities in the factors driving their biodiversity. Positive correlations between predicted invertebrate richness and litter organism richness may reflect

**a) Invertebrate Order Richness****b) Litter Organism Species Richness****c) Coleoptera Family Richness****d) Diptera Family Richness****e) Total Plant Species Richness****f) Native Plant Species Richness****g) Neophyte Plant Species Richness****h) Bird Species Richness**

(caption on next page)

**Fig. 4.** Heat maps that visually depict the conditional probabilities driving each model. These represent the strength of the relationships between each model's input parameters (patch area, vegetation height and connectivity current; values of bin ranges are shown in Table 4) and the predicted richness of each taxonomic group. Darker cells denote larger conditional probabilities, i.e. a higher likelihood of an outcome given that set of conditions, or a stronger relationship between the combination of input value and predicted richness represented by that cell in the heat map. *Note:* the homogenous appearance of high native plant richness is due to few data points being present in that range.

**Table 4**

Bin ranges for input parameter values in Fig. 4 heat maps of Bayesian model conditional probabilities.

Bins	Low	Moderately low	Moderate	Moderately high	High
Patch Area (ha)	0–1	1–5	5–10	10–100	100–500
Vegetation Height (m)	0–1	1–5	5–10	10–20	20–30
Connectivity Current	0–0.001	0.001–0.002	0.002–0.003	0.003–0.005	0.005–1

similarity or overlap in habitat dependency. Positive correlations were also present between Coleoptera and Diptera, and Diptera and bird richness (cf. Gagné & Fahrig, 2011).

Neophyte plant richness predictions tended to be weakly or negatively correlated to other groups, except total plant richness. This indicates that, at local scales, the planting or spread of non-native plants may occur at the expense of native plant species and to such an extent that neophytes threaten to dominate measures of overall plant richness. This contrasts with work done at larger spatial scales suggesting that non-native plants add to native plant richness to generate urban hot-spots of plant richness (Nielsen et al., 2014). Bird richness was also, and even more strongly, negatively related to total plant richness. Negative correlations between bird richness and multiple types of plant richness may reflect that avian richness is generally high in shrub and tree cover, but these habitat types tend to contain relatively few plant species.

The presence of low or negative correlations, varying sensitivities, disparate spatial patterns and taxon-specific conditional probabilities provide evidence that, whereas some taxa may respond similarly to landscape changes, trade-offs are likely to be present in how landscape variation impacts different taxa (Sushinsky, Rhodes, Possingham, Gill, & Fuller, 2013). Some variations in landscape-richness relationships are expected to reflect differences in the ecological effects of the selected landscape factors for the different taxa. The low and negative correlations between several invertebrate taxa and plant richness metrics, for example, are difficult to interpret in the context of the literature and may be an expression of nuance in biodiversity relationships that our research has not captured. Neophyte plants may be more common in highly managed gardens, and/or less supportive of biodiversity among other taxa than native plants (Manchester & Bullock, 2000; cf. Smith, Thompson, Hodgson, Warren, & Gaston, 2006); however this does not explain the low or negative correlations between various invertebrates and total or native plant richness predictions. Our results show that vegetation height plays a key role in predicting urban biodiversity, but that its relationship with overall richness summarises a complex range of relationships with subsets of the ecological community. Vegetation height may thus act here as a proxy for more specific aspects of vegetation structure or community composition, which in turn have variable impacts on ecological dynamics.

#### 4.3. Model performance and structure

Although the model error rates (Table 3) varied widely and were generally large, the mean rate among all models (60.5%) was comparable to results found in other studies applying BBNs to environmental systems (e.g. Aalders, 2008; Lemercier, Lacoste, Loum, & Walter, 2012; Taalab et al., 2015b). The highly complex and indirect nature of relationships between landscape structure and biodiversity fostered the expectation of a degree of uncertainty, stemming from sampling, system conceptualisation and natural variability (Regan, Colyvan, & Burgman, 2002). Nevertheless, the finding of comparable error rates, despite our model framework applying only three predictors

to such a complex question, supports the assertion that the predictors we consider here are strong determinants of urban biodiversity.

The way in which predictors were conceptualised and implemented here represents only one of many possible ways in which these characteristics can be quantified. Vegetation height was used as a proxy for maturity and vegetation structure to support applicability of findings to a wide audience of practitioners and researchers. However, other metrics that deal more directly with structural complexity are beginning to become more widely available with the increasing prevalence of LiDAR technology and approaches for processing the resulting data (Hancock et al., 2017). In the case of patch area, how best to define a habitat patch is not a new debate in landscape ecology (Kupfer, 2012). As implemented here, an area can be defined as a patch according to geometric rules that may vary in applicability between different taxa, and/or not necessarily reflect species responses to habitat patches. Choices of metric and measurement scale may be based on diverse criteria, but the methods introduced here appear effective.

An additional consideration, as with any modelling approach, is that models may be susceptible to mismatches between the spatial scale of inquiry and the functional scale of key drivers of biodiversity; moreover, our level of spatial precision (5 m) is unlikely to be available for most applications or studies in other urban areas. The analysis of models based on 5 m resolution data produced similar error rates to those based on 25 m data (see Supplementary Materials), even for the invertebrate taxa studied, suggesting that effectively capturing landscape-richness relationships is a more complex matter than simply adjusting the scale of inquiry. For instance, taxa such as birds may interact with their environment at coarser scales than those used here (e.g. Siriwardena, Calbrade, Vickery, & Sutherland, 2006), so may be better represented by considering landscape factors from a radius of 50 m or more around observation points (however, we note that the way connectivity was implemented helped to consider the surrounding landscape context). The differences in sensitivity analysis results between the 5 m and 25 m models (e.g. Coleoptera, Diptera and native plants) support the conviction that spatial biodiversity patterns are driven by different factors at different spatial scales (Pickett & Siriwardena, 2011). Potential scale mismatches between habitat use by taxa and spatial representation of the landscape are difficult to avoid; in practical terms (and for policy support), a standard, comparable scale was more useful for all taxa, rather than introducing greater uncertainty and impeding comparability through the use of taxon-specific scales.

A final scale-related consideration is that the prediction process embodied in the model presented here concerns point diversity, i.e. the expected diversity in a particular defined area, but does not capture beta-diversity across a site or habitat, which may also be an important consideration for biodiversity modelling (Ferrier, Manion, Elith, & Richardson, 2007) and for planning, where questions about biodiversity might typically be framed at the whole habitat or site level (e.g. a park). Solutions to this challenge are not straightforward; the reviewed data on which model structure is based come from samples at multiple scales concerning different taxa and landscape factors (Beninde et al., 2015)



**Table 5**  
Correlation coefficients (Rho) between model outputs at 25 m resolution. Calculated using Spearman's method. n = 273,303 in all cases.

	Invertebrate Order Richness	Litter Organism Species Richness	Coleoptera Family Richness	Diptera Family Richness	Total Plant Species Richness	Native Plant Species Richness	Neophyte Plant Species Richness
Litter Organism Sp. Richness	0.79	×					
Coleoptera Family Richness	0.29	0.24	×				
Diptera Family Richness	0.02	0.13	0.53	×			
Total Plant Species Richness	−0.18	−0.18	−0.21	−0.65	×		
Native Plant Species Richness	−0.27	−0.16	0.09	−0.18	0.40	×	
Neophyte Plant Species Richness	0.08	0.07	−0.34	−0.45	0.56	−0.25	×
Bird Species Richness	0.05	0.17	0.07	0.33	−0.44	−0.10	−0.47

and no single scale is likely to capture turnover effects adequately. Alternatives include explicit modelling of beta diversity (Ferrier et al., 2007) or approaches based on the summation of models for individual species, from which species composition (and hence both alpha and beta diversity) can be derived (Kattwinkel et al., 2009; Milanovich et al., 2012; Olden, Joy, & Death, 2006).

#### 4.4. The practical value and implications of BBN modelling of biodiversity

The challenge of moving from individual small-scale studies of urban biodiversity to general principles for understanding biodiversity and informing planning and conservation decisions at a city-wide scale is considerable. There is a need to explore novel potential approaches to developing such tools, and understand the situations in which they may be of practical value. The BBN modelling approach that we employ here has a number of important elements that can increase its potential (cf. Addison et al., 2013; Chee et al., 2016). First, it uses a generalised understanding of the key influences on biodiversity that can be based on empirical meta-analysis (as is the case in this example) or a more qualitative approach, such as expert opinion, where appropriate. The general framework can be adapted to make use of the highest quality information available and to take the local context into account through expert knowledge. A second advantage is the flexibility in creating both the structure and conditional probabilities of the network for a particular element of biodiversity. Finally, the BBN approach has the capacity to apply the influence network spatially, taking GIS data layers as input and deriving spatial biodiversity predictions for entire landscapes. The capacity to do this begins to make it possible to model biodiversity consequences of changing urban land use and form, which is a key element of taking account of biodiversity in the planning cycle (Norton et al., 2016).

Although we have demonstrated the feasibility of using a GIS-linked BBN approach for studying urban biodiversity, there remains a need for future research to compare the performance of this method against alternative approaches, particularly those conventionally favoured in biodiversity studies. Bayesian approaches in general offer different advantages to more conventional modelling methods (Aguilera, Fernández, Fernández, Rumí, & Salmerón, 2011; Jellinek et al., 2014; Taalab, Corstanje, Mayr, Whelan, & Creamer, 2015a) but will benefit from further development of relevant tools to produce more comprehensive and user-friendly interfaces and result formats in order to give them the broadest possible relevance. Nonetheless, we believe our approach has potential value as a method of biodiversity prediction in complex landscapes, and could provide a suitable mechanism for incorporation into a biodiversity module within an ecosystem service modelling suite such as InVEST (Tallis et al., 2014).

#### 4.5. Conclusions

We have demonstrated the feasibility of using GIS-coupled Bayesian Belief Networks to model biodiversity at fine spatial scales in complex and heterogeneous urban landscapes across a range of taxonomic groups. Our models performed with error rates similar to BBN models used in other environmental contexts. Our approach provides useful information on the sensitivity of biodiversity to different landscape structural features, thus providing ecological understanding that is relevant to planning decisions and assessment of urban development impacts. Our findings support the conviction that, broadly, richness in urban areas is increased by (1) the presence of large habitat patches, (2) high landscape connectivity, (3) tall vegetation and, by extension, mature vegetation communities (analogous to “bigger, better, more joined up”; see Lawton et al., 2010). Of these, vegetation height emerged as exerting a particularly strong influence on biodiversity across a range of taxonomic groups, although detailed biodiversity relationships may depend on more specific elements of vegetation structure and complexity than were directly considered here. Specific

responses to these drivers and directionality of associations do, however, differ between taxa and to some extent across the range of scales that we use. The results further highlight the established importance of landscape structural diversity, while equally showcasing the complex nature of relationships between landscape and biodiversity.

## Acknowledgements

The authors are grateful for the contributions of the following researchers in compiling data on invertebrate richness: Edward Lim and Emily Gravestock for Coleoptera and invertebrate order-level sorting; Vicki Senior for Diptera sorting; Rachel Clark for invertebrate sorting; Mark Telfer for Coleoptera identification checks, Andrew Godfrey for Diptera identification checks, and multiple undergraduate project students for processing invertebrate samples. Pseudoscorpions were identified by attendees at the Field Studies Council Pseudoscorpions workshop (June 2016). Mark Evans and Mark Gregory conducted the bird surveys. This research (grant number NE/J015067/1) was conducted as part of the Fragments, Functions and Flows in Urban Ecosystem Services (F<sup>3</sup>UES) project as part of the larger Biodiversity and Ecosystem Service Sustainability (BESS) framework. BESS is a six-year programme (2011–2017) funded by the UK Natural Environment Research Council (NERC) and the Biotechnology and Biological Sciences Research Council (BBSRC) as part of the UK's Living with Environmental Change (LWEC) programme. This work presents the outcomes of independent research funded by NERC and the BESS programme, and the views expressed are those of the authors and not necessarily those of the BESS Directorate or NERC.

## Author contributions

R.C., D.R.G., P.H.W., K.L.E. and J.A.H. conceived and designed the research. D.R.G. and R.C. constructed the BBN models. B.A.N., G.M.S., O.L.P., K.E.P., M.M. and J.P.R. collected and processed data. D.R.G., R.C., P.H.W., K.L.E. and J.A.H. framed the paper and wrote the first draft. D.R.G., M.M. and J.Z. performed modelling and analysis. All authors contributed to writing the paper.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.landurbplan.2019.05.012>.

## References

- Aalders, I. (2008). Modeling land-use decision behavior with Bayesian belief networks. *Ecology and Society*, 13(1)[https://doi.org/10.1890/1059-0933\(2008\)13\[1\]<131:Modelingland-usedecisionbehaviorwithBayesianbeliefnetworks>2.0.CO;2](https://doi.org/10.1890/1059-0933(2008)13[1]<131:Modelingland-usedecisionbehaviorwithBayesianbeliefnetworks>2.0.CO;2).
- Addison, P. F. E., Rumpff, L., Bau, S. S., Carey, J. M., Chee, Y. E., Jarrad, F. C., ... Burgman, M. A. (2013). Practical solutions for making models indispensable in conservation decision-making. *Diversity and Distributions*, 19(5–6), 490–502. <https://doi.org/10.1111/ddi.12054>.
- Aguilera, P. A., Fernández, A., Fernández, R., Rumí, R., & Salmerón, A. (2011). Bayesian networks in environmental modelling. *Environmental Modelling & Software*, 26(12), 1376–1388. <https://doi.org/10.1016/j.envsoft.2011.06.004>.
- Aitkenhead, M. J., & Aalders, I. H. (2009). Predicting land cover using GIS, Bayesian and evolutionary algorithm methods. *Journal of Environmental Management*, 90(1), 236–250. <https://doi.org/10.1016/j.jenvman.2007.09.010>.
- Anderson, B. J., Armsworth, P. R., Eigenbrod, F., Thomas, C. D., Gillings, S., Heinemeyer, A., ... Gaston, K. J. (2009). Spatial covariance between biodiversity and other ecosystem service priorities. *Journal of Applied Ecology*, 46(4), 888–896. <https://doi.org/10.1111/j.1365-2664.2009.01666.x>.
- Bagstad, K. J., Semmens, D. J., Waage, S., & Winthrop, R. (2013). A comparative assessment of decision-support tools for ecosystem services quantification and valuation. *Ecosystem Services*, 5, 27–39. <https://doi.org/10.1016/j.ecoser.2013.07.004>.
- Beninde, J., Veith, M., & Hochkirch, A. (2015). Biodiversity in cities needs space: A meta-analysis of factors determining intra-urban biodiversity variation. *Ecology Letters*, 18(6), 581–592. <https://doi.org/10.1111/ele.12427>.
- Bradbury, R. B., Hill, R. A., Mason, D. C., Hinsley, S. A., Wilson, J. D., Baltzer, H., ... Bellamy, P. E. (2005). Modelling relationships between birds and vegetation structure using airborne LiDAR data: A review with case studies from agricultural and woodland environments. *Ibis*, 147(3), 443–452. <https://doi.org/10.1111/j.1474-919X.2005.00438.x>.
- Casalegno, S., Anderson, K., Hancock, S., & Gaston, K. J. (2017). Improving models of urban greenspace: From vegetation surface cover to volumetric survey, using waveform laser scanning. *Methods in Ecology and Evolution*, 8(11), 1443–1452. <https://doi.org/10.1111/2041-210X.12794>.
- Chee, Y. E., Wilkinson, L., Nicholson, A. E., Quintana-Ascencio, P. F., Fauth, J. E., Hall, D., ... Rumpff, L. (2016). Modelling spatial and temporal changes with GIS and Spatial and Dynamic Bayesian Networks. *Environmental Modelling and Software*, 82, 108–120. <https://doi.org/10.1016/j.envsoft.2016.04.012>.
- Cox, D. T. C., Inger, R., Hancock, S., Anderson, K., & Gaston, K. J. (2016). Movement of feeder-using songbirds: The influence of urban features. *Scientific Reports*, 6(November), 37669. <https://doi.org/10.1038/srep37669>.
- Crawley, M. J., Harvey, P. H., & Purvis, A. (1996). Comparative ecology of the native and Alien Floras of the British Isles. Retrieved from *Philosophical Transactions: Biological Sciences*, 351(1345), 1251–1259. <http://www.jstor.org/stable/56200>.
- Deák, B., Hüse, B., & Tóthmérész, B. (2016). Grassland vegetation in urban habitats – Testing ecological theories. *Tuexenia*, 36, 379–393. <https://doi.org/10.14471/2016.36.017>.
- Dickson, B. G., Albano, C. M., Gray, M. E., McClure, M. L., Theobald, D. M., Anantharaman, R., ... Schumaker, N. (2018). Circuit-theory applications to connectivity science and conservation. *Conservation Biology*, 1–11. <https://doi.org/10.1111/cobi.13230>.
- Dyer, R. J., Gillings, S., Pywell, R. F., Fox, R., Roy, D. B., & Oliver, T. H. (2017). Developing a biodiversity-based indicator for large-scale environmental assessment: A case study of proposed shale gas extraction sites in Britain. *Journal of Applied Ecology*, 54(3), 872–882. <https://doi.org/10.1111/1365-2664.12784>.
- Evans, K. L., Chamberlain, D. E., Hatchwell, B. J., Gregory, R. D., & Gaston, K. J. (2011). What makes an urban bird? *Global Change Biology*, 17(1), 32–44. <https://doi.org/10.1111/j.1365-2486.2010.02247.x>.
- Ferrier, S., Manion, G., Elith, J., & Richardson, K. (2007). Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions*, 13(3), 252–264. <https://doi.org/10.1111/j.1472-4642.2007.00341.x>.
- Fordyce, J. A., Gompert, Z., Forister, M. L., & Nice, C. C. (2011). A hierarchical bayesian approach to ecological count data: A flexible tool for ecologists. *PLoS One*, 6(11), e26785. <https://doi.org/10.1371/journal.pone.0026785>.
- Fraser, F. C., Todman, L. C., Corstanje, R., Deeks, L. K., Harris, J. A., Pawlett, M., ... Ritz, K. (2016). Distinct respiratory responses of soils to complex organic substrate are governed predominantly by soil architecture and its microbial community. *Soil Biology and Biochemistry*, 103, 493–501. <https://doi.org/10.1016/j.soilbio.2016.09.015>.
- Gagné, S. A., & Fahrig, L. (2011). Do birds and beetles show similar responses to urbanization? *Ecological Applications*, 21(6), 2297–2312. <https://doi.org/10.1890/09-1905.1>.
- Gillespie, T. W., de Goede, J., Aguilar, L., Jenerette, G. D., Fricker, G. A., Avolio, M. L., ... Pataki, D. E. (2017). Predicting tree species richness in urban forests. *Urban Ecosystems*, 20(4), 839–849. <https://doi.org/10.1007/s11252-016-0633-2>.
- Grafius, D. R., Corstanje, R., Siriwardena, G. M., Plummer, K. E., & Harris, J. A. (2017). A bird's eye view: Using circuit theory to study urban landscape connectivity for birds. *Landscape Ecology*, 32(9), 1771–1787. <https://doi.org/10.1007/s10980-017-0548-1>.
- Hancock, S., Anderson, K., Disney, M., & Gaston, K. J. (2017). Measurement of fine-spatial-resolution 3D vegetation structure with airborne waveform lidar: Calibration and validation with voxelised terrestrial lidar. *Remote Sensing of Environment*, 188, 37–50. <https://doi.org/10.1016/j.rse.2016.10.041>.
- Hinsley, S. A., Hill, R. A., Bellamy, P., Broughton, R. K., Harrison, N. M., Mackenzie, J. A., ... Ferns, P. N. (2009). Do highly modified landscapes favour generalists at the expense of specialists? An example using woodland birds. *Landscape Research*, 34(5), 509–526. <https://doi.org/10.1080/01426390903177276>.
- Hothorn, T., Hornik, K., van de Wiel, M. A., & Zeileis, A. (2006). A lego system for conditional inference. *The American Statistician*, 60(3), 257–263. <https://doi.org/10.1198/000313006X118430>.
- Ives, C. D., Lentini, P. E., Threlfall, C. G., Ikin, K., Shanahan, D. F., Garrard, G. E., ... Kendall, D. (2016). Cities are hotspots for threatened species. *Global Ecology and Biogeography*, 25(1), 117–126. <https://doi.org/10.1111/geb.12404>.
- Jellinek, S., Rumpff, L., Driscoll, D. A., Parris, K. M., & Wintle, B. A. (2014). Modelling the benefits of habitat restoration in socio-ecological systems. *Biological Conservation*, 169, 60–67. <https://doi.org/10.1016/j.biocon.2013.10.023>.
- Kattwinkel, M., Strauss, B., Biedermann, R., & Kleyer, M. (2009). Modelling multi-species response to landscape dynamics: Mosaic cycles support urban biodiversity. *Landscape Ecology*, 24(7), 929–941. <https://doi.org/10.1007/s10980-009-9371-7>.
- Korb, K. B., & Nicholson, A. E. (2010). *Bayesian artificial intelligence*. CRC Press.
- Kovacs, K., Polasky, S., Nelson, E., Keeler, B. L., Pennington, D., Plantinga, A. J., & Taff, S. J. (2013). Evaluating the return in ecosystem services from investment in public land acquisitions. *PLoS One*, 8(6), e22202. <https://doi.org/10.1371/journal.pone.0062202>.
- Kupfer, J. A. (2012). Landscape ecology and biogeography: Rethinking landscape metrics in a post-FRAGSTATS landscape. *Progress in Physical Geography*, 36(3), 400–420. <https://doi.org/10.1177/0309133312439594>.
- Landuyt, D., Broekx, S., D'hondt, R., Engelen, G., Aertsens, J., & Goethals, P. L. M. (2013). A review of Bayesian belief networks in ecosystem service modelling. *Environmental Modelling and Software*, 46, 1–11. <https://doi.org/10.1016/j.envsoft.2013.03.011>.
- Lawton, J. H., Brotherton, P. N. M., Brown, V. K., Elphick, C., Fitter, A. H., Forshaw, J., ... Wynne, G. R. (2010). *Making space for nature: A review of England's wildlife Sites and ecological network*. Report to Defra. York. Retrieved from <http://webarchive.nationalarchives.gov.uk/20130402151656/http://archive.defra.gov.uk/environment/biodiversity/index.htm>.
- Lemercier, B., Lacoste, M., Loum, M., & Walter, C. (2012). Extrapolation at regional scale

- of local soil knowledge using boosted classification trees: A two-step approach. *Geoderma*, 171–172, 75–84. <https://doi.org/10.1016/j.geoderma.2011.03.010>.
- Lewis, D. D. (1998). Naïve (Bayes) at forty: The independence assumption in information retrieval. In C. Nédellec, & C. Rouveil (Eds.). *Machine learning: ECML-98* (pp. 4–15). Berlin: Springer. <https://doi.org/10.1007/BFb0026666>.
- Manchester, S. J., & Bullock, J. M. (2000). The impacts of non-native species on UK biodiversity and the effectiveness of control. *Journal of Applied Ecology*, 37(5), 845–864. <https://doi.org/10.1046/j.1365-2664.2000.00538.x>.
- Massimino, D., Johnston, A., & Pearce-Higgins, J. W. (2015). The geographical range of British birds expands during 15 years of warming. *Bird Study*, 62(4), 523–534. <https://doi.org/10.1080/00063657.2015.1089835>.
- McGarigal, K., Cushman, S. A., & Ene, E. (2012). *Fragstats v4: Spatial pattern analysis program for categorical and continuous maps*. Produced by the authors at the University of Massachusetts, Amherst. Available at: Amherst: University of Massachusetts Landscape Ecology Lab. <http://www.umass.edu/landeco/research/fragstats/fragstats.html>.
- McRae, B. H., Dickson, B. G., Keitt, T. H., & Shah, V. B. (2008). Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*, 89(10), 2712–2724. <https://doi.org/10.1890/07-1861.1>.
- McRae, B. H., Shah, V. B., & Mohapatra, T. K. (2013). *Circuitscape 4 user guide*. The Nature Conservancy <http://www.circuitscape.org>.
- Milanovich, J. R., Peterman, W. E., Barrett, K., & Hopton, M. E. (2012). Do species distribution models predict species richness in urban and natural green spaces? A case study using amphibians. *Landscape and Urban Planning*, 107(4), 409–418. <https://doi.org/10.1016/j.landurbplan.2012.07.010>.
- Millennium Ecosystem Assessment (2005). *Ecosystems and human well-being: Synthesis*. Retrieved from Washington D.C.: Island Press. <https://www.millenniumassessment.org/en/index.html>.
- Milton Keynes Council (2015). *Find out more about Milton Keynes*. Retrieved September 4, 2015, from <http://www.milton-keynes.gov.uk/jobs-careers/find-out-more-about-milton-keynes>.
- Nielsen, A. B., van den Bosch, M., Maruthaveeran, S., & van den Bosch, C. K. (2014). Species richness in urban parks and its drivers: A review of empirical evidence. *Urban Ecosystems*, 17(1), 305–327. <https://doi.org/10.1007/s11252-013-0316-1>.
- Norsys (2016). *Netica*. Retrieved from Vancouver: Norsys Software Corp. [www.norsys.com](http://www.norsys.com).
- Norton, B. A., Evans, K. L., & Warren, P. H. (2016). Urban biodiversity and landscape ecology: Patterns, processes and planning. *Current Landscape Ecology Reports*, 1(4), 178–192. <https://doi.org/10.1007/s40823-016-0018-5>.
- Office for National Statistics (2013). *2011 census, key statistics for built up areas in England and Wales (report), vol. 2015*. London: United Kingdom Office for National Statistics. Retrieved from <http://webarchive.nationalarchives.gov.uk/20160105160709/http://ons.gov.uk/ons/rel/census/2011-census/key-statistics-for-built-up-areas-in-england-and-wales/index.html>.
- Olden, J. D., Joy, M. K., & Death, R. G. (2006). Rediscovering the species in community-wide predictive modeling. *Ecological Applications*, 16(4), 1449–1460. [https://doi.org/10.1890/1051-0761\(2006\)016\[1449:RTSICP\]2.0.CO;2](https://doi.org/10.1890/1051-0761(2006)016[1449:RTSICP]2.0.CO;2).
- Ordnance Survey (GB) (2017). *OS mastermap topography layer [FileGeoDatabase geospatial data]*. Retrieved from EDINA Digimap Ordnance Survey Service <http://digimap.edina.ac.uk>.
- Pakhomov, S., Buntrock, J., & Chute, C. (2006). Automating the assignment of diagnosis codes to patient encounters using example-based and machine learning techniques. *Journal of the American Medical Informatics Association*, 13(5), 516–525. <https://doi.org/10.1197/jamia.M2077.Introduction>.
- Pickett, S. R. A., & Siriwardena, G. M. (2011). The relationship between multi-scale habitat heterogeneity and farmland bird abundance. *Ecography*, 34(6), 955–969. <https://doi.org/10.1111/j.1600-0587.2011.06608.x>.
- R Development Core Team (2016). *R: A language and environment for statistical computing*. Retrieved from Vienna, Austria: R Foundation for Statistical Computing. <http://www.r-project.org/>.
- Regan, H. M., Colyvan, M., & Burgman, M. A. (2002). A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecological Applications*, 12(2), 618. <https://doi.org/10.2307/3060967>.
- Rodríguez, J. P., Brotons, L., Bustamante, J., & Seoane, J. (2007). The application of predictive modelling of species distribution to biodiversity conservation. *Diversity and Distributions*, 13(3), 243–251. <https://doi.org/10.1111/j.1472-4642.2007.00356.x>.
- Siriwardena, G. M., Calbrade, N. A., Vickery, J. A., & Sutherland, W. J. (2006). The effect of the spatial distribution of winter seed food resources on their use by farmland birds. *Journal of Applied Ecology*, 43(4), 628–639.
- Smith, R. I., Barton, D. N., Dick, J., Haines-Young, R., Madsen, A. L., Rusch, G. M., ... Vikström, S. (2018). Operationalising ecosystem service assessment in Bayesian Belief Networks: Experiences within the OpenNESS project. *Ecosystem Services*, 29(5), 452–464. <https://doi.org/10.1016/j.ecoser.2017.11.004>.
- Smith, R. M., Thompson, K., Hodgson, J. G., Warren, P. H., & Gaston, K. J. (2006). Urban domestic gardens (IX): Composition and richness of the vascular plant flora, and implications for native biodiversity. *Biological Conservation*, 129(3), 312–322. <https://doi.org/10.1016/j.biocon.2005.10.045>.
- Stassopoulou, A., Petrou, M., & Kittler, J. (1998). Application of a Bayesian network in a GIS based decision making system. *International Journal of Geographical Information Science*, 12(1), 23–46. <https://doi.org/10.1080/136588198241996>.
- Stewart, G. B., Higgins, J. P. T., Schünemann, H., & Meader, N. (2015). The use of bayesian networks to assess the quality of evidence from research synthesis: 1. *PLoS One*, 10(4), e0114497. <https://doi.org/10.1371/journal.pone.0114497>.
- Sushinsky, J. R., Rhodes, J. R., Possingham, H. P., Gill, T. K., & Fuller, R. A. (2013). How should we grow cities to minimize their biodiversity impacts? *Global Change Biology*, 19(2), 401–410. <https://doi.org/10.1111/gcb.12055>.
- Taalab, K. P., Corstanje, R., Mayr, T. M., Whelan, M. J., & Creamer, R. E. (2015a). The application of expert knowledge in Bayesian networks to predict soil bulk density at the landscape scale. *n/a-n/a European Journal of Soil Science* (September) <https://doi.org/10.1111/ejss.12282>.
- Taalab, K. P., Corstanje, R., Zawadzka, J., Mayr, T., Whelan, M. J., Hannam, J. A., & Creamer, R. (2015b). On the application of Bayesian networks in digital soil mapping. *Geoderma*, 259–260, 134–148. <https://doi.org/10.1016/j.geoderma.2015.05.014>.
- Tallis, H. T., Ricketts, T., Guerry, A. D., Wood, S. A., Sharp, R., Nelson, E., ... Vogl, A. L. (2014). *Integrated valuation of environmental services and tradeoffs (InVEST) 3.1.0 user's guide*. Retrieved from Stanford: Natural Capital Project. <http://www.naturalcapitalproject.org/InVEST.html>.
- Thompson, K., & McCarthy, M. A. (2008). Traits of British alien and native urban plants. *Journal of Ecology*, 96(5), 853–859. <https://doi.org/10.1111/j.1365-2745.2008.01383.x>.